

A Step-indexed Semantic Model of Types for the Call-by-Name Lambda Calculus

Benedikt Meurer
 Compilerbau und Softwareanalyse
 Universität Siegen
 D-57068 Siegen, Germany
 meurer@informatik.uni-siegen.de

Abstract

Step-indexed semantic models of types were proposed as an alternative to purely syntactic safety proofs using subject-reduction. Building upon the work by Appel and others, we introduce a generalized step-indexed model for the call-by-name lambda calculus. We also show how to prove type safety of general recursion in our call-by-name model.

1 Introduction

Until recently, the most common way to prove type safety was by a purely syntactic proof technique called subject-reduction, which was adapted from combinatory logic by Wright and Felleisen [12]. One shows that each step of computation preserves typability (preservation) and that typable states are safe (progress).

This is not the only way though. Type safety can also be proved with respect to a semantic model. The semantic approach used in this paper avoids formalizing syntactic type expressions. Instead, one defines types as sets of semantic values. Using a technique called *step-indexing*, one then relates terms to these semantic types, and proves that typability implies safety. Instead of formalizing syntactic typing judgements, one formulates typing lemmata and proves their soundness with respect to the semantic model.

Related work Appel *et al.* introduced step-indexed models in the context of foundational proof

carrying code [5]. While they were primarily interested in low-level languages, they also applied their technique to a pure call-by-value λ -calculus with recursive types [6]. Our work generalizes the framework by Appel *et al.* to call-by-name by generalizing ground substitutions to terms instead of just values.

Ahmed *et al.* successfully extended the step-indexed models introduced by Appel *et al.* to general references and impredicative polymorphism [3, 4]. Hritcu *et al.* further extended it to object types, subtyping and bounded quantified types [9, 10]. They also indirectly considered the call-by-name λ -calculus using its well-known encoding in the ζ -calculus [1], including an encoding of the fixed point combinator [8].

Outline In section 2 we present the syntax and small step semantics of the programming language. Section 3 introduces semantic types and typing lemmata for the simply typed λ -calculus, which is extended with recursive types in section 4, while section 5 considers general recursion in the simply typed λ -calculus.

2 The language and its small-step semantics

The language we consider in this paper is the pure λ -calculus extended with constants, the simplest functional language that exhibits run-time errors (closed terms that “go wrong”). Its syntax is shown in Figure 1. We write $a[x \mapsto b]$ for the (capture

Variables: x, y, z, \dots
 Constants: $c ::= 0 \mid 1 \mid \dots$
 Terms: $a, b ::= c \mid x \mid \lambda x. a \mid a b$

Figure 1: Basic syntax

avoiding) substitution of b for all unbound occurrences of x in a . A term v is a *value* if it is a constant c or a closed term of the form $\lambda x. a$.

$$\frac{}{(\lambda x. a) b \rightarrow a[x \mapsto b]} \quad \frac{a \rightarrow a'}{a b \rightarrow a' b}$$

Figure 2: Small-step semantics

The operational semantics, as shown in small-step style in Figure 2, is entirely conventional [11]. We write $a_0 \rightarrow^k a_k$ if there exists a sequence of k steps such that $a_0 \rightarrow a_1 \rightarrow \dots \rightarrow a_k$. We write $a \rightarrow^* b$ if $a \rightarrow^k b$ for some $k \geq 0$. We say that a is *safe for k steps* if for any sequence $a \rightarrow^j b$ of $j < k$ steps, either b is a value or there is some b' such that $b \rightarrow b'$. Note that any term is safe for 0 steps. A term a is called *safe* if it is safe for every $k \geq 0$.

3 Semantic types

In this section we construct the methods for proving that a given term is safe in the call-by-name λ -calculus, using a simplified type system without recursive types. The semantic approach taken here considers types as indexed sets of values rather than syntactic type expressions.

Definition 1. A *type* is a set τ of pairs $\langle k, v \rangle$ where $k \geq 0$ and v is a value, and where the set τ is such that, whenever $\langle k, v \rangle \in \tau$ and $0 \leq j \leq k$, then $\langle j, v \rangle \in \tau$. For any term a and type τ we write $a :_k \tau$ if a is closed, and if, whenever $a \rightarrow^j b$ for some irreducible term b and $j < k$, then $\langle k - j, b \rangle \in \tau$.

Intuitively, $a :_k \tau$ means that the closed term a behaves like an element of τ for k steps of computation. That is, k computation steps do not suffice to prove that a does not terminate with a value of type τ . Note that if $a :_k \tau$ and $0 \leq j \leq k$ then $a :_j \tau$. Also, for a value v and $k > 0$, the statements $v :_k \tau$ and $\langle k, v \rangle \in \tau$ are equivalent.

Definition 2. A *type environment* is a mapping from variables to types. An *environment* (or *ground substitution*) is a mapping from variables to terms. For any type environment Γ and environment γ we write $\gamma :_k \Gamma$ if $\text{dom}(\gamma) = \text{dom}(\Gamma)$ and $\gamma(x) :_k \Gamma(x)$ for every $x \in \text{dom}(\gamma)$. We write $\Gamma \models a :_k \tau$ if $\gamma(a) :_k \tau$ for every $\gamma :_k \Gamma$, where $\gamma(a)$ is the result of replacing the unbound variables in a with their terms under γ . We write $\Gamma \models a : \tau$ if $\Gamma \models a :_k \tau$ for every $k \geq 0$.

Note that $\Gamma \models a : \tau$ can be viewed as a three place relation that holds on the type environment Γ , the term a , and the type τ . Utilizing this typing relation we can express static typing rules, which operate on terms with unbound variables. But first we observe that the safety theorem, stating “typability implies safety”, is a direct consequence of definitions 1 and 2 now, whereas in a syntactic type theory it is at least tedious to prove.

Theorem 3. *If $\emptyset \models a : \tau$, then a is safe.*

We can now construct semantic types and appropriate typing lemmata to derive true judgements of the form $\Gamma \models a : \tau$. Figure 3 gives the types and

$$\begin{aligned} \perp &\equiv \emptyset \\ \top &\equiv \{\langle k, v \rangle \mid k \geq 0\} \\ \text{Nat} &\equiv \{\langle k, c \rangle \mid k \geq 0\} \\ \tau \rightarrow \tau' &\equiv \{\langle k, \lambda x. a \rangle \mid \forall j < k \forall b. \\ &\quad b :_j \tau \Rightarrow a[x \mapsto b] :_j \tau'\} \end{aligned}$$

Figure 3: Semantic types

Figure 4 gives the typing lemmata for the simply typed λ -calculus. The remainder of this section is devoted to proving the soundness of these lemmata.

$$\begin{aligned} \frac{}{\Gamma \models x : \Gamma(x)} \quad & \frac{}{\Gamma \models c : \text{Nat}} \\ \frac{\Gamma \models a : \tau \rightarrow \tau' \quad \Gamma \models b : \tau}{\Gamma \models a b : \tau'} \\ \frac{\Gamma[x \mapsto \tau] \models a : \tau'}{\Gamma \models \lambda x. a : \tau \rightarrow \tau'} \end{aligned}$$

Figure 4: Semantic typing lemmata

The lemma for variables, stating $\Gamma \models x : \Gamma(x)$, follows directly from the definition of \models . The fact that \mathbf{Nat} is a type, and $\Gamma \models c : \mathbf{Nat}$, both follow immediately from the definition of \mathbf{Nat} . We now consider the lemmata for applications and lambda terms. First we have the following lemma which follows immediately from the definition of \rightarrow .

Lemma 4. *If τ and τ' are types then $\tau \rightarrow \tau'$ is also a type.*

Proof. By definition of \rightarrow it is obvious that $\tau \rightarrow \tau'$ is closed under decreasing index. \square

Lemma 5. *If $a_1 :_k \tau \rightarrow \tau'$ and $a_2 :_k \tau$, then $(a_1 a_2) :_k \tau'$.*

Proof. Since $a_1 :_k \tau \rightarrow \tau'$ and $a_2 :_k \tau$ we have that both a_1 and a_2 are closed, and if a_1 generates an irreducible term in less than k steps, that term must be a lambda term. Hence, the application $a_1 a_2$ either reduces for k steps without any top-level β -reduction, or there must be a lambda term $\lambda x. b$ such that $a_1 a_2 \rightarrow^j (\lambda x. b) a_2$ for some $j < k$.

In the first case, we know that $a_1 a_2$ is closed, and does not generate an irreducible term in less than k steps, and hence $a_1 a_2 :_k \tau'$.

Otherwise we have $a_2 :_{k-(j+1)} \tau$ by closure under decreasing index, and $\langle k-j, \lambda x. b \rangle \in \tau \rightarrow \tau'$ by Definition 1. $b[x \mapsto a_2] :_{k-(j+1)} \tau'$ follows by definition of \rightarrow . But now we have $a_1 a_2 \rightarrow^{j+1} b[x \mapsto a_2]$ and $b[x \mapsto a_2] :_{k-(j+1)} \tau$, and we can conclude $a_1 a_2 :_k \tau'$. \square

Theorem 6 (Application). *Let Γ be a type environment, let a_1 and a_2 be (possibly open) terms, and let τ and τ' be types. If $\Gamma \models a_1 : \tau \rightarrow \tau'$ and $\Gamma \models a_2 : \tau$, then $\Gamma \models a_1 a_2 : \tau'$.*

Proof. By Lemma 5 we have $\gamma(a_1 a_2) :_k \tau'$ for every $k \geq 0$ and γ , whenever $\gamma :_k \Gamma$, $\gamma(a_1) :_k \tau \rightarrow \tau'$ and $\gamma(a_2) :_k \tau$. Hence, we conclude $\Gamma \models a_1 a_2 :_k \tau'$ (for every $k \geq 0$). \square

Theorem 7 (Abstraction). *Let Γ be a type environment, let τ and τ' be types, and let $\Gamma[x \mapsto \tau]$ be the type environment that is identical to Γ except that it maps x to τ . If $\Gamma[x \mapsto \tau] \models a : \tau'$, then $\Gamma \models \lambda x. a : \tau \rightarrow \tau'$.*

Proof. Let $k \geq 0$, b be a closed term with $b :_k \tau$, and γ be an environment such that $\gamma :_k \Gamma$. Then $\gamma[x \mapsto b] :_k \Gamma[x \mapsto \tau]$, and since $(\gamma[x \mapsto b])(a) :_k \tau'$

and b is closed, we also have $\gamma(a[x \mapsto b]) :_j \tau'$ and $b :_j \tau$ for every $j < k$. Then $\langle k, \gamma(\lambda x. a) \rangle \in \tau \rightarrow \tau'$, and since $\gamma(\lambda x. a)$ is obviously closed, we conclude $\Gamma \models \lambda x. a :_k \tau \rightarrow \tau'$ (for every $k \geq 0$). \square

4 Recursive types

Recursive types were one of the main motivations behind the model of Appel *et al.* [6], and their results apply here, therefore we do not go into much detail. Figure 5 shows the recursion type operator μ , which computes a candidate fixed point of a function F from types to types by repeatedly applying the function to \perp , and the two typing lemmata for recursive types.

$$\mu F \equiv \{ \langle k, v \rangle \mid \langle k, v \rangle \in F^{k+1}(\perp) \}$$

$$\frac{\Gamma \models a : F(\mu F)}{\Gamma \models a : \mu F} \qquad \frac{\Gamma \models a : \mu F}{\Gamma \models a : F(\mu F)}$$

Figure 5: Recursive types

We will show that the typing lemmata in Figure 5 hold in the case where F is *well founded*. This is achieved by proving $\mu F = F(\mu F)$ for every well founded F , essentially proving that our recursive types are actually equi-recursive types, in contrast to iso-recursive types where μF is only isomorphic to $F(\mu F)$ via *roll* and *unroll* constructs on terms [2, 7].

Definition 8. The *k-approximation* of an indexed set τ is the subset

$$[\tau]_k = \{ \langle j, v \rangle \mid j < k \wedge \langle j, v \rangle \in \tau \}$$

of its elements whose index is less than k .

Obviously $[\tau]_k$ is a type whenever τ is a type. We now define a notion of well founded functional. Intuitively, a recursive definition of a type τ is well founded if, in order to determine whether or not $a :_k \tau$, it suffices to show $b :_j \tau$ for all terms b and indices $j < k$.

Definition 9. A *well founded functional* is a function F from types to types such that

$$[F(\tau)]_{k+1} = [F([\tau]_k)]_{k+1}$$

for every type τ and every index $k \geq 0$.

Lemma 10. *For every well founded functional F and every $k \geq 0$ we have:*

1. μF is a type
2. $\lfloor \mu F \rfloor_k = \lfloor F(\mu F) \rfloor_k$

Theorem 11. *If F is a well founded functional, then $\mu F = F(\mu F)$.*

See the paper of Appel and McAllester [6] for the proof sketch.

5 General recursion

As mentioned by Appel *et al.* [6], step-indexed types can also be used to simplify the semantic treatment of the fixed point rule to type recursive functions in the simply typed λ -calculus (without recursive types). Using the generalized framework presented in section 3, we are able to provide a direct, semantic soundness proof of the fixed point rule, which avoids any use of semantic domains, term orders, or monotonicity.

$$\frac{}{\text{fix } a \rightarrow a(\text{fix } a)} \quad \frac{\Gamma \models a : \tau \rightarrow \tau}{\Gamma \models \text{fix } a : \tau}$$

Figure 6: General recursion

We consider the standard fixed point operator [11], written $\text{fix } a$, for the call-by-name lambda calculus. The small step rule and the new typing lemma is shown in Figure 6. The remainder of this section is devoted to proving the soundness of the semantic typing lemma.

Lemma 12. *If $a :_k \tau \rightarrow \tau$, then $(\text{fix } a) :_k \tau$.*

Proof. By induction on k . Since $a :_k \tau \rightarrow \tau$ implies $a :_j \tau \rightarrow \tau$ for every $j < k$, we also have $(\text{fix } a) :_j \tau$ for every $j < k$ by induction hypothesis, and using Lemma 5 we also have $(a(\text{fix } a)) :_j \tau$. Of course, $(\text{fix } a)$ is closed whenever a is closed. So assume that there is some irreducible term b and some $j < k$ such that $(\text{fix } a) \rightarrow^j b$. This implies $j > 0$ and $(\text{fix } a) \rightarrow a(\text{fix } a) \rightarrow^{j-1} b$, and since $(a(\text{fix } a)) :_{j-1} \tau$ we have $\langle k - j, b \rangle \in \tau$. Hence, we conclude $(\text{fix } a) :_k \tau$. \square

This leads immediately to the following theorem, stating the soundness of the typing lemma for the call-by-name fixed point operator as shown in Figure 6.

Theorem 13 (General recursion). *Let Γ be a type environment, let a be a term, and let τ be a type. If $\Gamma \models a : \tau \rightarrow \tau$, then $\Gamma \models \text{fix } a : \tau$.*

6 Conclusion

We have presented a step-indexed model for the call-by-name lambda calculus, and used it to prove the safety of a type system with recursive types. We also proved safety of general recursion in our framework.

References

- [1] M. Abadi and L. Cardelli. *A Theory of Objects*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1st edition, 1996.
- [2] M. Abadi and M. P. Fiore. Syntactic Considerations on Recursive Types. In *Proceedings of the 11th Annual IEEE Symposium on Logic in Computer Science, LICS '96*, pages 242–252, Washington, DC, USA, 1996. IEEE Computer Society.
- [3] A. J. Ahmed. *Semantics of types for mutable state*. PhD thesis, Princeton University, Princeton, NJ, USA, 2004.
- [4] A. J. Ahmed, A. W. Appel, and R. Virga. A Stratified Semantics of General References Embeddable in Higher-Order Logic. In *Proceedings of the 17th Annual IEEE Symposium on Logic in Computer Science, LICS '02*, pages 75–86, Washington, DC, USA, 2002. IEEE Computer Society.
- [5] A. W. Appel and A. P. Felty. A semantic model of types and machine instructions for proof-carrying code. In *Proceedings of the 27th ACM SIGPLAN-SIGACT symposium on Principles of programming languages, POPL '00*, pages 243–253, New York, NY, USA, 2000. ACM.

- [6] A. W. Appel and D. McAllester. An indexed model of recursive types for foundational proof-carrying code. *ACM Trans. Program. Lang. Syst.*, 23:657–683, September 2001.
- [7] K. Crary, R. Harper, and S. Puri. What is a recursive module? *SIGPLAN Not.*, 34:50–63, May 1999.
- [8] K. Fisher, F. Honsell, and J. C. Mitchell. A lambda calculus of objects and method specialization. *Nordic J. of Computing*, 1:3–37, March 1994.
- [9] C. Hritcu. *A Step-indexed Semantic Model of Types for the Functional Object Calculus*. M.Sc. thesis, Programming Systems Lab, Department of Informatics, Saarland University, 2007.
- [10] C. Hritcu and J. Schwinghammer. A Step-indexed Semantics of Imperative Objects. *Logical Methods in Computer Science*, 5(4):1–48, December 2009.
- [11] B. C. Pierce. *Types and Programming Languages*. MIT Press, Cambridge, MA, USA, 2002.
- [12] A. K. Wright and M. Felleisen. A syntactic approach to type soundness. *Inf. Comput.*, 115:38–94, November 1994.